

Towards a Praxis of Critical Digital Sport History

In 1989, a year before the founder of the modern internet Sir Tim Berners-Lee authored the code that became the first web browser and four years before the release of the first graphical web browser, Orville Vernon Burton published an article entitled “History’s Electric Future” in the *Organization of American Historians’ Newsletter*.¹ “Just as the weight-driven mechanical clock is the metaphor for the early modern period and the steam engine represents the industrial revolution,” he wrote, “the computer is the symbol of our age.” Our age, in Burton’s depiction, was a world where researchers could have instant access to the entire catalogue of libraries; modems allowed scholars to query online databases, mainframe computers, and other researcher’s “microcomputers” to both investigate their holdings as well as undertake simple computational tasks. They could augment their own records by transferring files from these remote digital locations to their own microcomputers. Libraries, he foretold, would “convert books and journals into machine-readable text” allowing simultaneous access by users from far-flung places. Researchers would be able to take affordable hand-held scanners into the archives to transform paper materials into machine-readable text. “Electronic books, serials, maps and documents with sound and music, pictures, color and electronic painting, movement and graphics will all make the next generation of historical works more exciting and instructional than at present.” Full-text databases would allow the organization, indexing,

¹ O. V. Burton, “History’s Electric Future,” *OAH (Organization of American Historians) Newsletter* 17, no. 4 (1989): 12–13.

and retrieval of information at one's fingertips. Researchers would communicate with one another via message boards that would allow them to discover common interests, share materials, and ultimately "produce a more sophisticated history." That history facilitated by computers, offered scholars the "rare opportunity to change the scale and magnitude of historical inquiry and ultimately to affect the questions historians ask."

Burton rightly identified the computer as a tool within the historian's craft. It would transform the pace and products of historical scholarship allowing primary source materials to be considered hand-in-hand with a scholar's secondary argumentation. The computer would enable side-by-side comparisons and could provide a virtual environment for scholars to debate the merits of evidence openly and without prolonged delays via listservs, forums, and later social media. Identifying library and archival materials, analyzing and augmenting sources, and producing narrative machine-readable texts would facilitate scholarship that Burton believed would be more accessible, less expensive, and ultimately more pleasing to the reader or viewer. But it would also challenge scholars in ways that were unanticipated. Historians would not just have to learn how to use these new computers and technologies like file transfer protocols and command lines. They would also have to recognize their lack of understanding about the computer and its associated tasks---design, programming, networking, etc.—at a more global scale. For every affordance the personal computer could offer as many problems and limitations would be introduced to the practice of research.²

² For example, Burton projected that storage would become so inexpensive as to render it possible for scholars to engage with the entire human record. While he was correct in that the costs of storage would decline dramatically over decades, the deluge of digital materials introduced significant challenges related to both organizing these materials as well as their long term preservation.

The near simultaneous release of low-cost personal computers and the dramatic investment in internet technologies within the United States created “History’s Electric Future” as one of expansive opportunity in the 1990s. Universities and libraries invested in server farms, internet infrastructures, and the purchase of computers and mainframes for its faculty’s use. Large-scale technology groups were created to manage the developing computer frameworks that would blanket campuses. So impactful would this investment in the internet be, that scholars telling the history of historian’s work with computers regularly start their stories not with the 1940s and 1950s, but with the early 1990s.³ The contributions of Harriet and Frank Owsley, Merle Curti, William O. Ayedelotte, and others who transformed manuscript records into quantitative data that could be tabulated and sorted via IBM-owned Hollerith machines are not related as digital history.⁴ Nor does the work of other social historians, who utilized mainframe computers from the 1940s through the late 1980s, receive much attention when historians seek to explore the roots of “digital

³ Stephen, “The Differences between Digital History and Digital Humanities,” *Dr Stephen Robertson*, May 23, 2014,

<http://drstephenrobertson.com/blog-post/the-differences-between-digital-history-and-digital-humanities/>.

⁴ In the late 1940s, Curti and others were using statistical analyses to transform individual records into datasets that allowed them to better understand the census and its role in topics as varied as the “Old South”, the American frontier, and the British Parliament. They spent hours creating hard-copy datasets that turned manuscript records into individually-prepared machine-readable records that could be run through a computer. Given the rudimentary nature of computer development at this time, each “job” or analytical operation had to be configured separately from one another. These “runs” often took hours and even days to complete. This led not just to voluminous stacks of punch cards that were subject to frequent revision but also to the requirement that scholars who sought to use these machines be trained in the computer’s language and ways in which the machine relied on mathematical operations to complete its work. There is tremendous potential for an enterprising scholar to recover the work of these scholars as contributors to the history of computing. It is important for us to consider these scholars as part of the “digital” history subfield not just to recognize their relationship to the long history of computing but to place historians within the histories of humanities computing that are being written and dominated by those in literary and linguistic computing. More simply, the work of historians, particularly social historians, who utilized mainframe computers from the 1940s through the late 1980s deserves attention and study in their own right as we explore what it means to be a historian and what the craft of the historian is and has been.

history.” These scholars disappear historiographically because they are “quantitative” historians rather than “digital” historians.

Instead the narrative of innovation and early adoption begins with Edward Ayers’s *The Valley of Shadow* (1993) and Roy Rosenzweig’s *Who Built America?* (1994), which remain widely recognized as “the first” historical projects to take advantage of hypertextual and multimedia-driven narrative, respectively. Ayers began *Valley* in 1991 as a comparative monograph of the North and South. Yet, through a partnership with the Institute for Advanced Technology in the Humanities at the University of Virginia, *Valley* was broadened “as a research library in a box, enabling students at places without a large archive to do the same kind of research as a professional historian.” It was one of two demonstration projects that allowed the University of Virginia to illustrate the potential of information technology in humanities fields to IBM.⁵ Ayers focused on the utilization of network technologies to produce computer-enabled archives coupled with interactive narrative that would allow him to tell a multiplicity of stories rather than the singular story that was encouraged by the structure of a monograph. Likewise, Rosenzweig developed his co-authored work to illustrate the potential of the cd-rom to deliver supplemental materials to teachers that would accompany print materials included within monographs. Unlike the quantitative historians of the 1940s and 1950s where the digital work was situated as illustration in the forms of charts or as evidence in the forms of percentages and

⁵ As scholars critique the role of digital history as complicit in both the neo-liberal university and the economics of research, the role of IBM in soliciting and supporting early digital history work should receive significant attention. On the digital humanities and neoliberalism, see Richard GrusinR. Grusin, “The Dark Side of Digital Humanities: Dispatches from Two Recent MLA Conventions,” *Differences* 25, no. 1 (January 1, 2014): 79–92, doi:10.1215/10407391-2420009; Miriam Posner, “What Alt-Ac Can Do, and What It Can’t,” *Miriam Posner’s Blog*, November 25, 2013, <http://miriamposner.com/blog/what-alt-ac-can-do-and-what-it-cant/>.

graphs, the work of Ayers and Rosenzweig sought to bring primary documents, in their original form, to their audiences, which included teachers, researchers, as well as visitors to cultural sites.

The early pioneers of computationally-assisted history have also been situated outside of the longer origins of “humanities computing”, the forerunner to today’s digital humanities.⁶ Susan Hockney noted in her history of humanities computing that historians supposedly found the technological limitations of computing too cumbersome to tolerate:

Most large-scale datasets were stored on magnetic tape, which can only be the other and so software was designed to minimize the amount of tape movement. Random access to data such as happens on a disk was not possible. Data had therefore to be stored in a serial fashion. This was not so problematic for textual data, but for historical material it could mean the simplification of data, which represented several aspects of one object (forming several tables in relational database technology), into a single linear stream. This in itself was enough to deter historians from embarking on computer-based projects.⁷

The origins of humanities computing has been overly deterministic in drawing its history not from the wide variety of innovators and early adopters from fields as diverse as history,

⁶ Patrik Svensson, “Humanities Computing as Digital Humanities” 3, no. 3 (2009), <http://digitalhumanities.org/dhq/vol/3/3/000065/000065.html>; Patrik Svensson, “The Landscape of Digital Humanities” 4, no. 1 (Summer 2010), <http://www.digitalhumanities.org/dhq/vol/4/1/000080/000080.html>; John M. Unsworth, “The State of Digital Humanities, 2010” (Digital Humanities Summer Institute, Victoria, BC, Canada, June 7, 2010), www3.isrl.illinois.edu/~unsworth/state.of.dh.DHSI.pdf; Susan Schreibman, Ray Siemens, and John Unsworth, *Companion to Digital Humanities (Blackwell Companions to Literature and Culture)*, Hardcover, Blackwell Companions to Literature and Culture (Oxford: Blackwell Publishing Professional, 2004), <http://www.digitalhumanities.org/companion/>.

⁷ Susan Hockey, “The History of Humanities Computing,” in *A Companion to Digital Humanities* (Oxford: Blackwell, 2004), 1–19, <http://onlinelibrary.wiley.com/doi/10.1002/9780470999875.ch1/summary>.

media studies, arts, and sociology, but instead from a literary-centric trajectory that emphasizes the work of scholars engaged in the work of creating concordances, exploring computational linguistics, and later digital editions enabled by the creation of standardized text markup for use on the internet.⁸ The consequence of this for humanities computing, particularly in the US context, was one of perceived delayed adoption by those outside of the literary and literature fields. Rather than integrating the work of Curti and others with Father Roberto Busa and his efforts, historians have largely been ignored within the continuum of scholars working with computers.

The consequences of historians being situated outside of humanities computing are clear. First, digital history is spoken of as if it is a relatively recent field of study where there is little previous scholarship to draw upon. That which is recognized is most frequently public history work presented via internet technologies. Secondly, historians who work with computers as tools of critical analysis and critique, “digital historians” working from the Ayers–Rosenzweig school of public history emphasis, and those who look to digital technology and the internet as a site of critique rarely engage with one another. These three paths (analysis, presentation, and critique) may blend. But those who use computational methods for analysis (usually emphasized as quantitative historians) are

⁸ The narrative of humanities computing is generally as follows: Father Busa and his work in the automated analysis of written texts for the purposes of creating a concordance in the 1940s opens most narrative trajectories of humanities computing. Humanities computing is then narrated through 1960s efforts to assist in compiling dictionaries, into scholarly organizations of the 1970s and 1980s, which promoted text analysis, authorship attribution, and electronic editions, where it finally merges in the 1990s with the work of Ayers and others for a brief moment in that they both leverage the internet as one of their primary forms of information delivery. Yet, this history of humanities computing elides scholars, particularly historians, who were invested in the creation of large machine-readable datasets in addition to those who were engaged in early work in computer-assisted cataloging, archiving, and information processing. On Father Busa, see Thomas Nelson Winter, “Roberta Busa, S.J., and the Invention of the Machine-Generated Concordance,” *Classical Bulletin*, 75.1 (1999), 6 and

largely separate from those who act as consumers and producers of material for current digital platforms; they too are divorced from those who treat computers and the internet as a site of historical study. The history of computing and technology, computational history, and digital history parallel one another rather than intersect. “Digital history” is the default domain of those who are producing history via internet-driven technologies rather than those who might be using digital technologies for analytical purposes or studying the history of computing technology. In part, this separation is a function of methodological training: learning about the history of computing and technology is separate from training in computational methods and digital history methods. Students may take a quantitative methods course or a digital history course or a history of computing course but it is a rare student who would learn all three. And even rarer is a student who has the pleasure of a computational methods course for humanists where they explore statistics, simulation and modeling for 2d and 3d environments, as well as author code to develop their own computational algorithms. Digital history courses tend to function as a broad introductory survey or as an exploration of a single methodological approach combined with a technological expertise (for example, digitization combined with text encoding methods or working with geographic information systems or quantitative methods using R or SPSS). On the one hand, this history-centric training offers us well-trained historians; on the other, it means that historians are oft unprepared to face the interdisciplinary critique their work might engender. Nor are they belatedly encouraged to confront the technology itself as a site of inquiry. And, more problematically, historians working digitally might assume a measure of “innovativeness” or “uniqueness” that is not merited when their work is

evaluated by non-historians.

This constraint of disciplinarity is not just a function of methodological training but also one of academic structuralism. The overwhelming majority of scholars working today in digital history focus on internet-based technologies and publication avenues that are supported by university-sponsored offices of technology (e.g. wordpress, omeka, or social media platforms) and digital technologies made available for free for public audiences (e.g. omeka, scalar, historypin, google maps). Even the choice of which software will be made available at a university-level might impact the results of a historian's work (e.g. ESRI versus QGIS, SPSS versus R, Gephi versus NodeXL, etc.). Systems of peer-review, tenure and promotion, and even publication have yet to confront this troubling trend that rewards claims of innovation and public access without regard for interdisciplinary assessment.⁹ The American Historical Association provides guidance to scholars and history departments in their "Guidelines for the Evaluation of Digital Scholarship in History".¹⁰ They define digital history as "scholarship that is either produced using computational tools and methods or presented using digital technologies." However, digital history---be it computational research, via internet technologies and production, or as a site of critique---exists as a space that intersects with a multiplicity of disciplines. The digital is both the domain and the medium of many including, but not limited to: engineering,

⁹ For example, library and archival professional's contribution to digital history work has largely been ignored as they are not historians. We see this most clearly as well with public practitioners who might act as consumers of digital history but are not allowed to participate in systems of promotion and review without a Ph.D. or other terminal credential.

¹⁰ American Historical Association, "Guidelines for the Evaluation of Digital Scholarship in History," accessed April 30, 2016, <https://www.historians.org/teaching-and-learning/digital-history-resources/evaluation-of-digital-scholarship-in-history/guidelines-for-the-evaluation-of-digital-scholarship-in-history>.

computing, library and information science, informatics, graphics and design, the humanities, social sciences, etc. As such, as we transition to talking about digital sport history, we must begin by not just understanding digital history as something that exists along a historical continuum of scholars and researchers who have been producing scholarship since the 1940s, but also as interdisciplinary at its core because of its very engagement with the digital which is the domain of many. The American Historical Association elides this interdisciplinary history of historian's computing and research methods when they note that professional evaluation should be conducted by their disciplinary [read: historian] colleagues. Even more problematically, they also degrade the role of technologists and librarians from peers and contributors to digital history projects to trainers of faculty for the purposes of evaluation: "Most colleges and universities have staff in place whose job it is to monitor and promote new technologies. Librarians, in particular, have long been involved in professional conversations regarding new technologies of teaching and scholarship. Many of them will be delighted to hold workshops and address faculty in groups or as individuals."¹¹ I point to these standards to suggest an important tension that Gary Osmond and Murray G. Phillips recognize in their conclusion to *Sport History in the Digital Era*, "digital history disrupts traditional historical practice in a multitude of ways: it provides viable alternatives to the culture of individually-driven scholarship, to the ability to work from a single-discipline perspective,

¹¹ Ibid. The AHA does slightly ameliorate their position later when they recommend that "since digital scholarship often includes collaborations, departments should consider developing protocols for evaluating collaborative work, such as co-authored works, undergraduate research, crowdsourcing, and development of tools."

and to the practice of producing linear narratives.”¹² Yet, the focus on disruption which often utilizes similar language to that of innovation obscures the continuum upon which historians continue to work even after the digital turn.

Historians engage with primary sources, actively construct relationships between primary and secondary sources, and author theories of the past backed by evidence. Traditional monographs with structured narratives used the disciplinary convention of citation to chart the evidentiary basis of argumentation. The sources, or in computer parlance “data” of one’s argument appeared piecemeal in charts, graphs, or footnotes. Prior to the advent of digital primary sources, scholars had to travel to the archive to pull sources to confirm accuracy. With Burton’s electronic future, scholars could begin to augment their monograph with full-text primary sources. And in the years since, large digital repositories of primary sources have enabled the real-time comparison and vetting of primary materials by scholars that Ayers and others envisioned being used by teachers. We’ve transferred our disciplinary practice of citation and evidence from analog to digital. Today’s challenges in using tweets or blogs are, in large measure, similar to the concerns of earlier evidentiary insertions. Questions of privacy, anonymity, authority and the like were confronted by oral historians who sought to bring those “new” forms of evidence to mainstream history in the same way that digital historians have brought forum posts, tweets, blogs, and reddit as new types of evidence for scholarly consideration.

As digital historians celebrate the liberation of sources from physical form (be they from analog to digital surrogates or in born-digital form), historians though have largely

¹² Gary Osmond and Murray G. Phillips, eds., *Sport History in the Digital Era* (Urbana ; Chicago ; Springfield: University of Illinois Press, 2015).

ignored the responsibility associated with born-digital and digitized data as an object of scholarly research. The AHA and others can see the merits of the digital archive as an act of authorship by a historian; but they largely ignore data outside of a searchable, visual archive as a product of historical scholarship. The result is that most historian's first act of digital history is to craft their own digital archive based on a topical or chronological era. It is not to contribute new data to a pre-existing digital archive. Here, we should recognize the work of information scientists who have demonstrated the value of research data both for potential contribution and reuse by other scholars but also as a commitment to transparent research methods. "The curation of research data – raw and abstracted material created as part of research processes and which may be used again as the input to further research – carries with it the burden of capturing and preserving not only the data itself, but information about the methods by which it was produced."¹³

The form and expression of data (or in historian's parlance "sources") are consistent only in that they are consistently shifting. Just as scholars once had to reconcile the voices of marginalized peoples that rose to the front of our discipline in the 1970s, so too must today's historian reconcile the often anonymous and disembodied voices of the internet age and the ways in which digital technologies and computers are determining what is being written and preserved since 1990. It is important that we maintain our recognition that power, preservation, and analysis are key to our understanding of both the existence

¹³ "An Introduction to Humanities Data Curation," *DH Curation Guide*, March 28, 2014, <https://guide.dhcuration.org/contents/intro/>; C. L Palmer, A. H Renear, and M. H Cragin, "Purposeful Curation: Research and Education for a Future with Working Data," *Status: Published or Submitted for Publication*, 2008; Dorothea Salo, "Digital Curation," Syllabus, *Library and Information Science: LIS 855, Digital Curation*, (2010), <http://scientopia.org/blogs/bookoftrogool/2011/01/18/syllabi-and-how-rapidly-they-become-obsolete/>.

and the meaning of these sources. Just as we are taught to “read” images, moving pictures, music, manuscripts, newspapers, and physical objects, so too must we also “read” digital objects, networks, and projects in their manifest forms as sites of production and contextualization.

Following in the tradition of speculative design, a few illustrations from sport history might assist in understanding how we must both read objects and their context and, just as importantly, note any activities or operations upon the object both as they are created and in their end form. In 1993, Michael Oriard published *Reading Football: How the Popular Press Created an American Spectacle*.¹⁴ An exploration of football as a cultural text, Oriard utilized microfilm of nineteenth century newspapers to conduct a close-reading that would allow him to consider how newspapers and popular periodicals constructed football as both game and as cultural object. To complete his research, Oriard spent hours pouring through microfilm to identify sports-pages, taking longhand notes, and duplicating articles of particular interest. In his introduction, he notes the transformative nature that microfilm afforded him: providing both access to these sources but also notes their poor quality limited his abilities in specific ways. His ultimate product was a monograph that includes quotation, illustration, and narration. Today, we might imagine a multitude of digital research processes that a contemporary scholar could take to study nineteenth century newspapers that could support, modify, or outright dispute Oriard’s conclusions about the history of football and its meaning in the popular press.

Scholar one might elect to reproduce Oriard’s primary source research using an

¹⁴ Michael Oriard, *Reading Football: How the Popular Press Created an American Spectacle*, Cultural Studies of the United States (Chapel Hill: University of North Carolina Press, 1993).

online search engine or library catalogue. This scholar might review each footnote, identify a corresponding digital surrogate if one is available, download the article, and then redisplay it in an online repository or website as a companion to Oriard's work where they might annotate or dispute Oriard's conclusions. The benefits to this approach would be that it would allow a tight assessment of Oriard's conclusions: when looking at the entirety of his primary sources (and secondary for that matter) would scholar one reach the same conclusion? While it might be tempting to conclude that scholar one is only one step divorced from the physical newspaper, scholar one is in fact at least three steps into the research process: from analog to digital surrogate via digitization and metadata cataloguing, through searching via keywords or metadata fields, and into a stage of visual display that may or may not be accessible to the scholar based on their institutional affiliation. The technological options for this scholar could manifest in a number of ways: 1) a website site where they post each individual article, summarize Oriard's conclusions and pose their own; 2) an Omeka digital exhibit that leverages Dublin core metadata to describe each article which they could then use to craft narrational relationships between them via the exhibit feature; or 3) use Hypothes.is, a plug in that leverages annotation to enable sentence-level critique or note-taking on top of any web-based object. The decision regarding which technology platform the scholar selects will effect everything from what the scholar will need to contribute (or customize), to how the objects will appear to the user, to the long term sustainability of the project for researchers who might want to explore scholar one's work decades down the line.

Scholar two might elect to use an application programming interface to batch

download the entire *Chronicling America* repository as well as any other digital newspaper repositories, and then conduct their own study of identified articles relating to football. Scholar two might be less interested in making the primary sources available to other researchers than analyzing these sources for the scholar's own purposes. Importantly, for scholar two, just because there is no act of "making" or "creating" associated with their approach does not mean the research process is any less potentially complex. *Chronicling America* requires the identification of keywords to create search results; the possibilities of false-positive results would require either individual review or the continual revision of keywords. Articles would also have to be assessed for issues of duplication. This would be in addition to issues of digital remediation and data curation already assessed against the primary sources. As Martin Johnes and Bob Nicholson noted in "Sport History and Digital Archives in Practice", optical character recognition, search algorithms, and information disambiguation in addition to digital remediation can complicate using newspapers within digital history research.¹⁵ Scholar two would need to note for others which sources were included in these databases, what search algorithms and keywords were used, how the dataset might have been later resolved in relation to issues of false positives, etc. In the case of our first scholar, his selection methodology is duplicative of Oriard. Thus, he'd need to investigate how Oriard originally put together his research methodology. How did he choose his newspapers and periodicals? Could scholar one identify errors or gaps that newly available material might reveal? In the case of our second scholar, though, the selection methodology derives not just from their own selection principles but also

¹⁵ Martin Johnes and Bob Nicholson, "Sport History and Digital Archives in Practice" in Osmond and Phillips, *Sport History in the Digital Era*.

constraints of the technologies the source has inhabited. Poor optical character recognition might elide articles that pertain to the scholar's search query. Problematic keywords or metadata might construct a different view of the search results for each scholar visiting the interface. "The newspapers in the Library of Congress's *Chronicling America* collection divide each issue into pages, not logical articles, and the *Making of America* magazine collection specifies only the page on which each article begins."¹⁶ As a result, scholar two might not generate a complete article return via their searching making close-reading like Oriard completed problematic. And, much more significantly, scholar two needs to recognize that *Chronicling America* is not a full single digital repository but rather an aggregator of state-based digitization efforts of newspapers. Because it is grant-funded, some states are not represented while others are overrepresented. This might lead scholar two to have issues of under and overrepresentation in the corpus. Northern urban sportwriters might discuss football much differently from those along the edges of rural Mississippi. Yet, if it is not clearly noted by the scholar in their exploration of the *Chronicling* interface that Mississippi papers are not included, the scholar might reach erroneous conclusions about the pervasiveness of popular media nationally. Additionally, newspapers are added regularly to the aggregator which means that later scholars might hold scholar two accountable for conclusions about Mississippi because scholar two did not note Mississippi's exclusion. This type of research process and methodological reveal must be done not just in the process of data collection but also as scholar two modifies the data

¹⁶ "Computational Methods for Uncovering Reprinted Texts in Antebellum Newspapers – Viral Texts," accessed May 1, 2016, <http://viraltexts.org/2015/05/22/computational-methods-for-uncovering-reprinted-texts-in-antebellum-news-papers/>.

as grouping, dividing, subdividing, and discarding occur.

Rather than beginning from an attempt to reconstruct Oriard's archive through digital means and close-reading, a third scholar might seek to understand the circulation of football teams in alignment with what popular press had to say about their play more generally. Let's assume that a cache of newspapers were made available both as images but also as plain-text files that were complete and pristine---every sports article written in the nineteenth century in whole in clear machine-readable text. Using current methods in text analysis borrowed from linguistic computing, scholar three could analyze the plain text files to identify language associated with sentiment. Using the Python programming language coupled with the Natural Language Processing Toolkit, scholar three could attempt to classify every article written about football for attitudes, opinions, and emotions. This method would allow the scholar to group articles together based on criteria like positive or negative feelings as well as consider subjective adjectives that might signal bias or conflict.¹⁷ This type of classification might reveal that particular genres of sportswriters offer a higher rate of racialized language or that specific newspapers are more or less likely to publish divisive language in its articles. Prior to loading the newspaper data into the classification and sentiment analysis platform, the scholar would need to process and "clean" the data to ensure that it does not report false results.¹⁸ The

¹⁷ Rachele Sprugnoli et al., "Towards Sentiment Analysis for Historical Texts," *Digital Scholarship in the Humanities*, July 27, 2015, fqv027, doi:10.1093/llc/fqv027.

¹⁸ Luciano Barbosa and Junlan Feng, "Robust Sentiment Detection on Twitter from Biased and Noisy Data," in *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, COLING '10 (Stroudsburg, PA, USA: Association for Computational Linguistics, 2010), 36–44, <http://dl.acm.org/citation.cfm?id=1944566.1944571>; T. Horton et al., "Quite Right, Dear, and Interesting": Seeking the Sentimental in Nineteenth Century American Fiction," *Digital Humanities*, 2006; D. J. Newman and S. Block, "Probabilistic Topic Decomposition of an Eighteenth-Century American Newspaper," *Journal of the American Society for Information Science and Technology* 57, no. 6 (2006): 753–767.

work of “cleaning” data is itself an act of mediation that requires both content knowledge as well as a discerning eye towards how the technology will interact with the data. And, in the case of sentiment analysis, the underlying training algorithms that trigger notations of sentiment might not have been aligned to nineteenth century newspapers. Instead, the algorithmic basis might be contemporary restaurant reviews, twitter, or even Netflix movie reviews. Thus, scholar three might find that the results of algorithmic analysis are inconsequential or too noisy for their purposes. This might necessitate creating a new training corpus or partnering with a computational linguist who can assist in steering scholar three towards other sentiment analysis softwares. Scholar three will find that experimentation will form a core activity.

Today’s savvy digital historian must, like those who constantly query the role of archives and archiving as an act of power, consider technology and digital platforms as both sites of possibility and of peril. As historians, we must recognize the digital, and its multiplicity of forms, as historical objects that are produced, interpreted, and contested. It is not sufficient to simply leverage technology as a platform to promote our work or to contribute to a global digital repository. Instead, digital historians must encounter the computer and the digital with a skeptical eye to understand that technology exists within its own frame of production and as a framework through which new knowledge is produced. Computers and the networks they comprise are not black box apparatus that should be simply used by scholars. We must actively work to understand how decisions about design, programming, expression, interface, networking, access, sustainability, etc. produce and privilege certain types of history. These are the same queries we leveraged

about archives in the 1970s and 1980s that we must now offer of the computer, the internet, and the digital.

We must hold our scholarship not up to the standards of digital sport history, which is relatively recent, but instead to the interdisciplinary fields that are actively working in the specific methodologies and with the technologies that we might engage with as researchers, teachers, and practitioners. The general desire to engage with the trend of making or consuming digital objects, placing scholarship on the web, or creating projects without attention to the conditions of their production will ultimately limit the potential of digital sport history. “Criticality” marks the potential of digital sport history rather than just the use of digital tools and platforms to explore how sport is being discussed in digital forums, via digital social media platforms, and as a topic of archival production. Further, by moving towards a critical digital sport history we can begin to reward not just the final products of historical scholarship but also recognize the work that comprises the various stages of historical research and interdisciplinary methodology. Gary Osmond and Murray G. Phillips in their conclusion to *Sport History in the Digital Era* recognize this potential: “digital history is a collaborative venture, multidisciplinary at its core, not wedded to linear narratives, created with community involvement in mind, and employs the multimedia capacities of digital technology to create histories whose shape, form, and function are different from professionally sanctioned historical products: monographs, journal articles, and thesis.”¹⁹ For Osmond and Phillips, “digital history has the capacity to challenge, unsettle, and reinvent the fundamental principles of history making.” Significantly, though,

¹⁹ Osmond and Phillips, *Sport History in the Digital Era*.

as of yet, the practice of *doing* digital history has not matched this theorization of digital history as somehow innately collaborative, multidisciplinary, community-driven, and non-linear.

As sport history is relatively recent to the field of digital history and humanities computing, we have a tremendous opportunity to shape the future of our scholarship by learning from the long history of digital history, humanities computing, and digital humanities. That future should recognize the very real affective impact our entrance into both digital history and digital humanities can have. Bethany Nowviskie, in her timely critique of the digital humanities, wrote that “our newer colleagues, who are most visible online, make two assumptions: they think that all of this [digital humanities] is new and they think that the current scene is all there is.”²⁰ Calling this the “eternal September”, Nowviskie highlights a fundamental tension between established scholars and those who are new to their exploration. The work of bringing others into the field can be exciting and yet incredibly arduous. New scholars can bring additional disciplinary and subdisciplinary concerns, updated or revised methods, and a questioning nature that can highlight assumptions and boundaries that have shaped ongoing digital scholarship. Their curiosity can reinvigorate areas of study and communities of scholars. Yet, the “Eternal September” can also lead new scholars to foolish conclusions about the status of the field, the types of scholarship that have come before, and the conventions that interdisciplinarity might demand of them. Doing one’s interdisciplinary homework about standards, methods, and secondary scholarship can be daunting but it is required when we recognize the long

²⁰ “Eternal September of the Digital Humanities « Bethany Nowviskie,” accessed October 20, 2010, <http://nowviskie.org/2010/eternal-september-of-the-digital-humanities/>.

trajectory of digital history and sport history's attempt to engage with it.

As digital sport history considers a robust future, we should learn from the expansion of digital history which was driven by individual projects spread across a wide thematic, chronological, and technological spectrum. We should encourage a evaluative framework that does not just privilege the results of digital sport history scholarship but also the values that we elect to hold. Lisa Spiro wrote in her pivotal essay "This is Why We Fight: Defining Values of the Digital Humanities of five core values: openness, collaboration, collegiality and connectedness, diversity, and experimentation. These values should form the cornerstone of our ongoing development with digital sport history rather than an attempt to enforce a singular methodology or vision of what constitutes digital sport history. More simply, we should recognize our products as aligned to interdisciplinary research processes that, while they may hold primary and secondary sources and the act of narration a key, will also seek to critically interrogate the digital as object, site, and medium of production. To work as a scholar in digital sports history, then, we must simultaneously engage sport history, scholarship from humanities computing/digital humanities, and the disciplines that engage with the appropriate methodology and technologies to the type of research we seek (e.g. statistics, media studies, etc.). And, we must commit to doing so as scholars who recognize what these other disciplines have to offer us in terms of the challenges and possibilities of working in and with digital technologies and computation. The time to place something on the internet (via a website, blog, tweet, repository, etc) without consideration of how the technology functions as well as issues of access, labor, sustainability, ethics, etc. is long past. We should look to disciplines beyond our own to

explore the approaches, methods, and products of digital scholarship and computational research and the lessons they can offer us. Doing so will allow us to integrate sport history more seamlessly with ongoing work in library and information science, archival science, computer science, media studies, etc. And, it would demonstrate sport history's future as one of critical, intentional engagement with interdisciplinary research rather than a faddish adoption.